



A measurement policy in stochastic linear filtering problems

C. Bruni, G. Koch, F. Papa*

Dipartimento di Informatica e Sistemistica "A. Ruberti", Sapienza Università di Roma, Via Ariosto 25, 00185 Roma, Italy

ARTICLE INFO

Article history:

Received 28 July 2010

Received in revised form 25 November 2010

Accepted 30 November 2010

Keywords:

Linear systems filtering

Measurement cost

Estimate accuracy

ABSTRACT

With reference to a (discrete time) linear system filtering problem, we consider the problem of online deciding at which instant we actually measure and process the output values. We propose a convenient measurement policy, able to sensibly reduce the measurement cost, while keeping the estimate accuracy at satisfactory levels. By some probabilistic arguments we analyse the upper bounds for the measurement and non-measurement time intervals.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The need for investigating the deep relationship between communication and control theory has been significantly acknowledged in the pertinent literature since the corner stone contribution of Norbert Wiener [1].

More recent contributions in the past decade are found in [2–5].

A peculiar example of interaction between communication and control in a dynamical system refers to the output measurement process, which should provide information about the state of the controlled system and, at the same time, be the object of a suitable policy aimed toward the saving of measurement cost.

The measurement policy problem might be addressed to various issues with reference to the use of the information carried by the measurement itself. For the case in which the estimate variance is minimized we quote [6]. For the case in which the reference problem is the control of a dynamical system, we mention [7,8] and the more recent [9–13].

In this work we assume that measurement processing is addressed to the estimation of the state of a discrete time linear system (linear filtering problem). The reference estimation procedure in this context is the Kalman filter, which is widely used in many applications, and in particular in biomedical signal processing [14–17].

The problem we formulate here is, at each time instant, to decide whether to measure (and process) the output values.

Measuring would indeed increase the information content and improve the accuracy of the state estimate. At the same time it involves a cost, which might be evaluated not only from an economical point of view, but also from the aspect of potential damage of the measured “object”, such as in biomedical applications.

In order to be sensible, the decision rule has to only rely on the information available before the measurement is actually performed.

More specifically, we propose to acquire and process the output any time the foreseen information lost in the previous non-measurement times exceeds a given threshold related to the measurement cost itself.

The procedure is explained in Section 2, where the information lost is evaluated by comparing the one step forecast of the state and the foreseen contribution of the innovation process.

Sections 3 and 4 investigate the expected length of the measurement and non-measurement time intervals.

* Corresponding author.

E-mail addresses: brunic@dis.uniroma1.it (C. Bruni), san_lorenzo@libero.it (G. Koch), papa@dis.uniroma1.it, fedepapa1983@gmail.com (F. Papa).

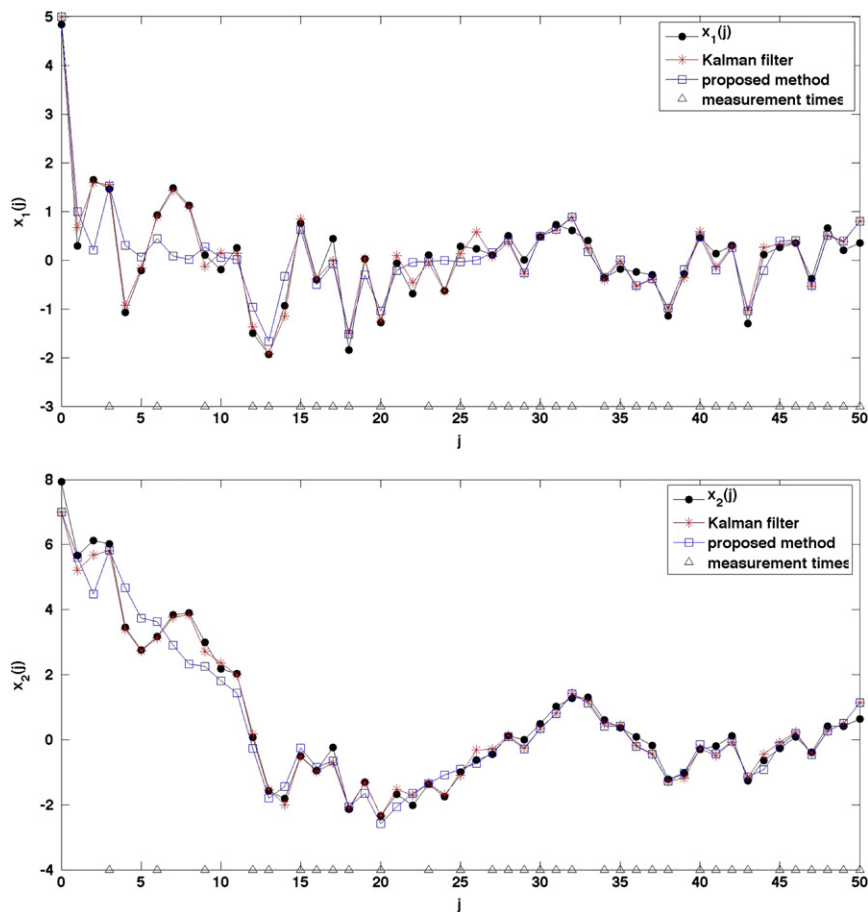


Fig. 1. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

Finally, in Section 5 the procedure is experimented with reference to a stable and to an unstable time-invariant linear system, and compared to the classical Kalman filter. In particular the saving of measurement along with the information lost are investigated by varying the value of the threshold, and of the input and output noise variances.

The examples show that a significant saving in the measurement cost might be achieved even at the expense of a loss in accuracy.

2. Kalman filtering: whether or not to measure the output process

Let us consider a linear discrete time (finite dimensional) dynamical system:

$$x(j+1) = A(j)x(j) + B(j)u(j), \quad j = 0, 1, \dots \quad (2.1)$$

where $x(j) \in \mathbb{R}^n$, $\forall j$, is the state of the system at time j , while the input $\{u(j) \in \mathbb{R}^p, j = 0, 1, \dots\}$ is assumed to be a Gaussian, zero mean, white noise process, with given covariance matrix $\Psi_u(j) = E\{u(j)u^T(j)\}$. We also assume the initial state $x(0)$ to be a Gaussian vector, independent of the process u , with $\bar{x}(0) = E\{x(0)\}$, $\Psi_x(0) = E\{x(0)x^T(0)\}$.

Let us consider the output process:

$$z(j) = C(j)x(j) + v(j), \quad j = 1, 2, \dots \quad (2.2)$$

with $z(j) \in \mathbb{R}^q$, $\forall j$, and the output noise $\{v(j) \in \mathbb{R}^q, j = 1, 2, \dots\}$ is assumed to be a Gaussian, zero mean, white noise process, independent of $x(0)$ and u , with given covariance matrix $\Psi_v(j) = E\{v(j)v^T(j)\}$.

As already mentioned in the introduction, in this paper we intend to discuss whether and when to actually measure the output, trying to reach a trade-off between information content and measurement cost.

Therefore, in the following, for any given process $f(j)$, we denote by $\hat{f}(j)$ the best (conditioned minimum square error) estimate of $f(j)$ given all past *actually measured* output values up to time j .

As is well known, $\hat{x}(j)$ is iteratively provided as a solution of the Kalman filter equation. Indeed, for any j such that $z(j+1)$ is actually measured, this equation is written as:

$$\hat{x}(j+1) = A(j)\hat{x}(j) + \alpha(j+1), \quad (2.3)$$

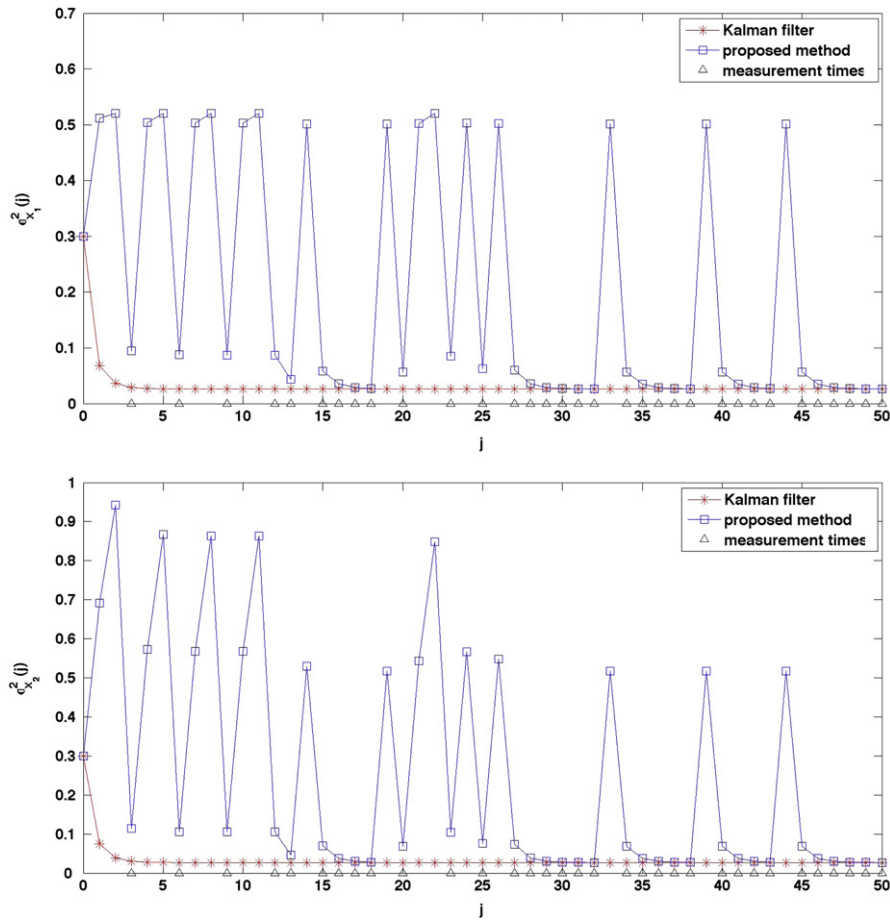


Fig. 2. Evolution of covariance matrix elements $\sigma_{x_1}^2(j)$ and $\sigma_{x_2}^2(j)$.

with

$$\alpha(j+1) = F(j+1)\rho(j+1), \quad (2.4)$$

$$\rho(j+1) = z(j+1) - C(j+1)A(j)\hat{x}(j), \quad (2.5)$$

$$F(j+1) = \Psi_e(j+1)C^T(j+1)\Psi_v^{-1}(j+1), \quad (2.6)$$

and $\Psi_e(j+1)$ is the estimate error covariance matrix which solves the Riccati equation:

$$\Psi_e(j+1) = \{[A(j)\Psi_e(j)A^T(j) + B(j)\Psi_u(j)B^T(j)]^{-1} + C^T(j+1)\Psi_v^{-1}(j+1)C(j+1)\}^{-1}. \quad (2.7)$$

In case $z(j+1)$ is skipped, the (forecast) equation is written as:

$$\hat{x}(j+1) = A(j)\hat{x}(j), \quad (2.8)$$

and the estimate error covariance matrix is given by:

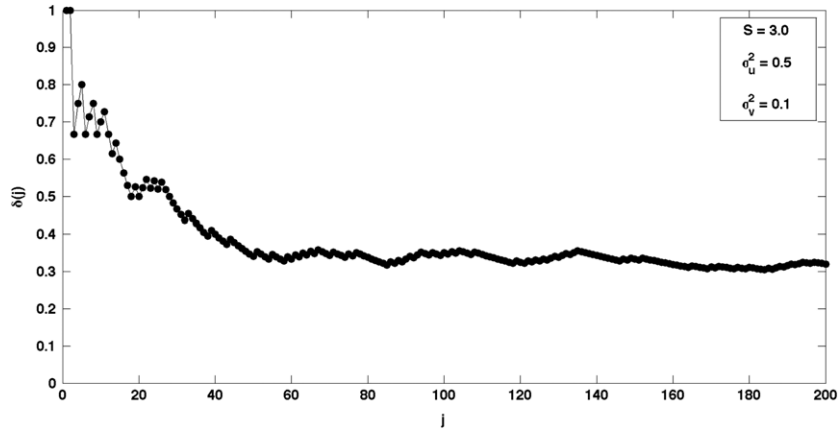
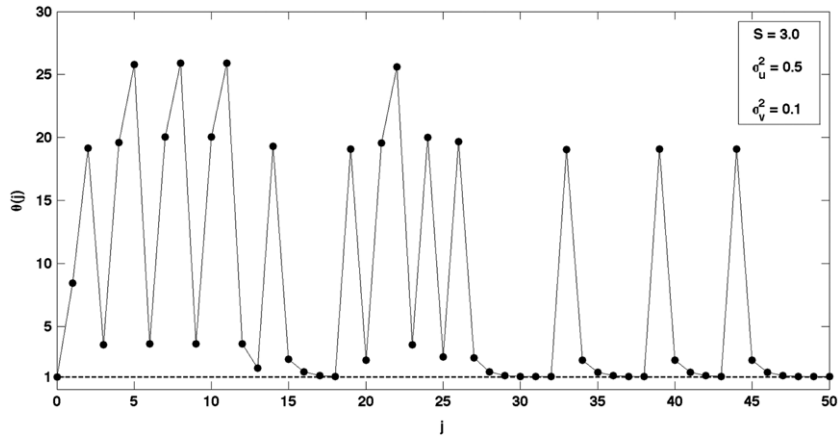
$$\Psi_e(j+1) = A(j)\Psi_e(j)A^T(j) + B(j)\Psi_u(j)B^T(j). \quad (2.9)$$

The previous equations may be formally unified, as j runs, in the form:

$$\begin{cases} \hat{x}(j+1) = A(j)\hat{x}(j) + \tilde{F}(j+1)[z(j+1) - \tilde{C}(j+1)A(j)\hat{x}(j)], & j = 0, 1, \dots, \\ \hat{x}(0) = \bar{x}(0), \end{cases} \quad (2.10)$$

and

$$\begin{cases} \Psi_e(j+1) = \{[A(j)\Psi_e(j)A^T(j) + B(j)\Psi_u(j)B^T(j)]^{-1} + \tilde{C}^T(j+1)\Psi_v^{-1}(j+1)\tilde{C}(j+1)\}^{-1}, \\ j = 0, 1, \dots, \quad \Psi_e(0) = \Psi_x(0), \end{cases} \quad (2.11)$$

(a) Measurement saving $-\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio $-\theta(j) = \frac{\text{Tr}\{\Psi_{\tilde{e}}(j)\}}{\text{Tr}\{\Psi_{\tilde{e}}^{(K)}(j)\}}$.**Fig. 3.** Effectiveness of the proposed procedure.

where

$$\tilde{F}(j+1) = \Psi_{\tilde{e}}(j+1)\tilde{C}^T(j+1)\Psi_v^{-1}(j+1), \quad (2.12)$$

and

$$\tilde{C}(j+1) = \begin{cases} C(j+1) & \text{in case } z(j+1) \text{ is actually measured,} \\ 0 & \text{otherwise.} \end{cases} \quad (2.13)$$

To actually measure the output process at a given time j , we propose it to be decided on the basis of a comparison between the expected loss of information over the current non-measurement time interval and the cost that the measurement itself at time j would imply.

More precisely, for any time $i > 0$, we assess the relative expected loss of information corresponding to the non-measurement of $z(i)$ by the ratio $\chi(i)$:

$$\chi(i) = \frac{E_{i-1}\{\tilde{x}^T(i)\tilde{x}(i)\}}{\tilde{x}^T(i)\tilde{x}(i)}. \quad (2.14)$$

In (2.14), $\tilde{x}(i)$ is the best estimate of $x(i)$ given all actually measured output values up to $i-1$, as well as $z(i)$; $E_{i-1}\{\cdot\}$ is the mean value given all actually measured output values up to $i-1$; $\tilde{x}(i) = E_{i-1}\{x(i)\} = A(i-1)\hat{x}(i-1)$. We just remark that, for any i , $\chi(i)$ is a non-negative quantity. An easy computation shows that:

$$\begin{aligned} \chi(i) &= \frac{E_{i-1}\{\tilde{x}^T(i)\tilde{x}(i)\}}{\tilde{x}^T(i)\tilde{x}(i)} = \frac{E_{i-1}\{[\tilde{x}(i) + \alpha(i)]^T[\tilde{x}(i) + \alpha(i)]\}}{\tilde{x}^T(i)\tilde{x}(i)} \\ &= \frac{\|\tilde{x}(i)\|^2 + \text{Tr}\{\Psi_{\alpha}(i)\}}{\|\tilde{x}(i)\|^2} = 1 + \frac{\text{Tr}\{\Psi_{\alpha}(i)\}}{\|\tilde{x}(i)\|^2}. \end{aligned} \quad (2.15)$$

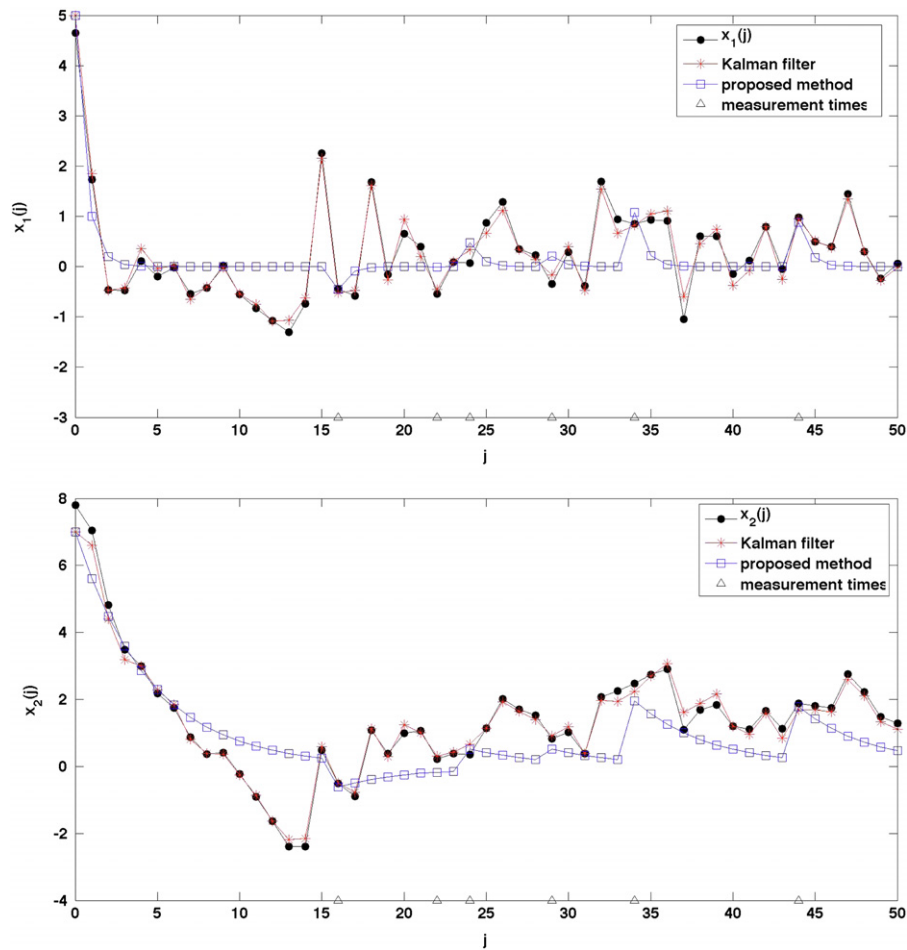


Fig. 4. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

In (2.15), $\Psi_\alpha(i) = E_{i-1}\{\alpha(i)\alpha^T(i)\}$ is the covariance matrix (conditioned upon all past actually measured output values up to time $i - 1$) of $\alpha(i)$. Note that $\alpha(i)$ is independent of $\tilde{x}(i)$, given all measurements up to time $i - 1$. Note also that the value of $\Psi_\alpha(i)$ can be achieved as the difference between (2.9) and (2.7) for $j = i - 1$.

In case $i = 1$, $\chi(1)$ in (2.14) is evaluated with reference to the available information about the initial distribution of $x(0)$.

Let h, j be any two time instants, with $j > h$. Then we define the total expected loss of information ratio over the time interval $h + 1, h + 2, \dots, j$ as:

$$\omega(j, h) = \sum_{i=h+1}^j \chi(i) = \sum_{i=h+1}^j \left[1 + \frac{\text{Tr}\{\Psi_\alpha(i)\}}{\|\tilde{x}(i)\|^2} \right] = j - h + \sum_{i=h+1}^j \frac{\text{Tr}\{\Psi_\alpha(i)\}}{\|\tilde{x}(i)\|^2}. \quad (2.16)$$

We are now able to formulate and apply the following decision rule.

Decision rule 2.1. For any time j , let k be the last actual measurement time before j . Then we measure (and process) $z(j)$ if and only if:

$$\omega(j, k) > S(j), \quad (2.17)$$

where $S(j)$ is a suitable threshold related to the measurement cost at time j .

Clearly, for any j such that $S(j) \leq 1$ the whole decision problem stated here loses interest, since (2.17) necessarily holds and the measurement of $z(j)$ has to be made.

We observe that, at each time j , the computational burden implied by Decision rule 2.1 simply amounts to evaluating the last term $\chi(j)$ in the summation (2.16). This in turn is essentially equivalent to the j th step in the classical Kalman filtering. In other words, the procedure we suggest does not increase the computational task, while saving (if this is the case) the measurement cost.

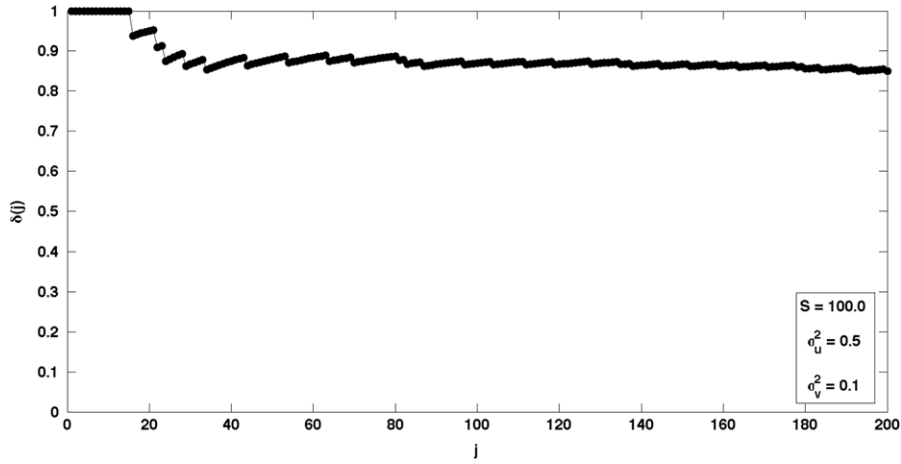
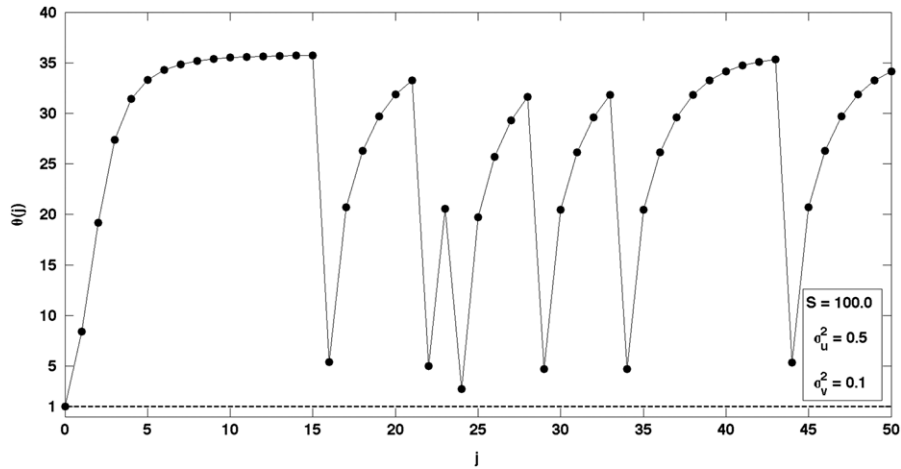
(a) Measurement saving $-\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio $-\theta(j) = \frac{\text{Tr}[\psi_e(j)]}{\text{Tr}[\psi_e^{(k)}(j)]}$.

Fig. 5. Effectiveness of the proposed procedure.

3. How long to keep on measuring the output process?

Let k be a time in which we performed an output process measurement. In the previous section we proposed a criterion about how to decide whether to keep measuring z , based on the function $\omega(j, k)$, $j = k + 1, \dots$, as defined by (2.16).

The question now arises about how long the measuring interval T would be, where $T + k$ is the first non-measurement time after k . Of course, at a fixed time k , we shall only be able to give T a probabilistic characterization.

A first result is the following.

Theorem 3.1. Under the current assumptions and notations, denoting by $P_k(\cdot)$ the probability distribution conditioned with respect to all actually available measurements up to time k and by $E_k\{\cdot\}$ the corresponding mean value, we have that:

$$P_k(T > 1) = I(\{\omega(k+1, k) > S(k+1)\}), \quad (3.1)$$

$$P_k(T > r) \leq P_k(T > 1) \prod_{j=k+1}^{k+r-1} \gamma(j+1), \quad r = 2, 3, \dots, \quad (3.2)$$

and

$$E_k\{T\} \leq 1 + P_k(T > 1) \left(1 + \sum_{r=2}^{\infty} \prod_{j=k+1}^{k+r-1} \gamma(j+1) \right), \quad (3.3)$$

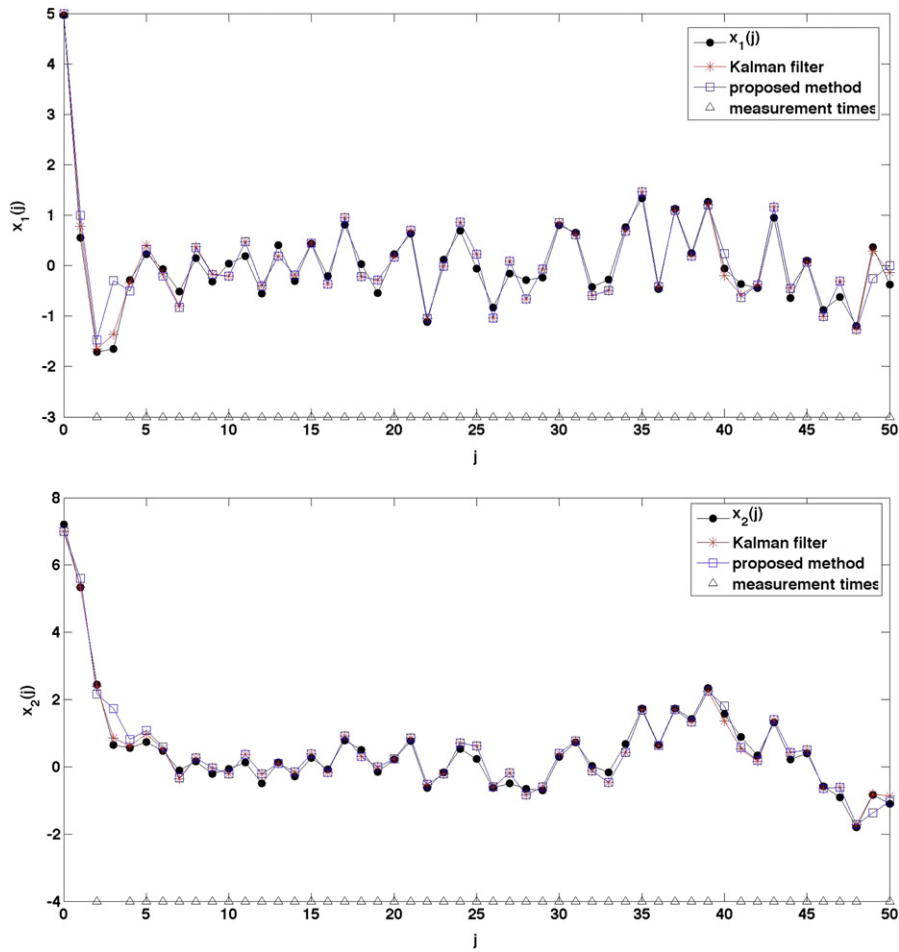


Fig. 6. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

where $I(\mathcal{E})$ is the indicator function of event \mathcal{E} and:

$$\gamma(j+1) = P_k(\|A(j)F(j)\rho(j)\|^2 \leq s(j+1)), \quad (3.4)$$

$$s(j+1) = \frac{\text{Tr}\{\Psi_\alpha(j+1)\}}{S(j+1)-1}. \quad (3.5)$$

Proof. We first note that (3.1) immediately follows from the very definition of T , according to which, given all actually available measurements up to time k , $T = 1$ if and only if $\omega(k+1, k) \leq S(k+1)$.

Next, we observe that:

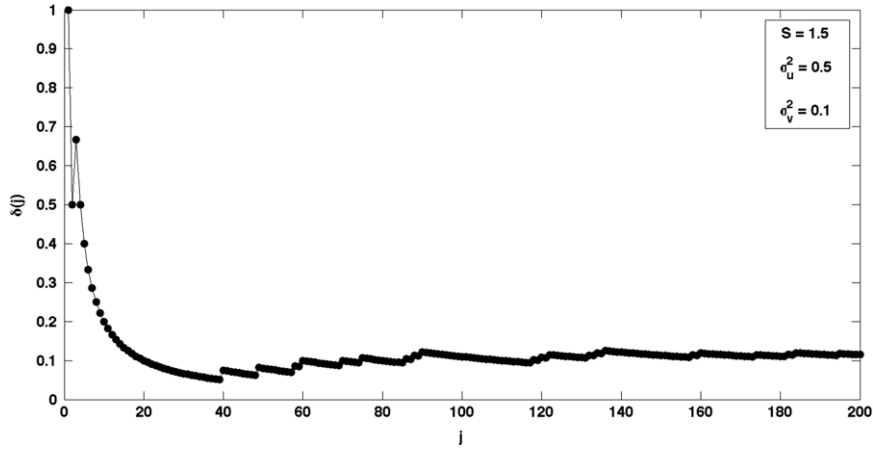
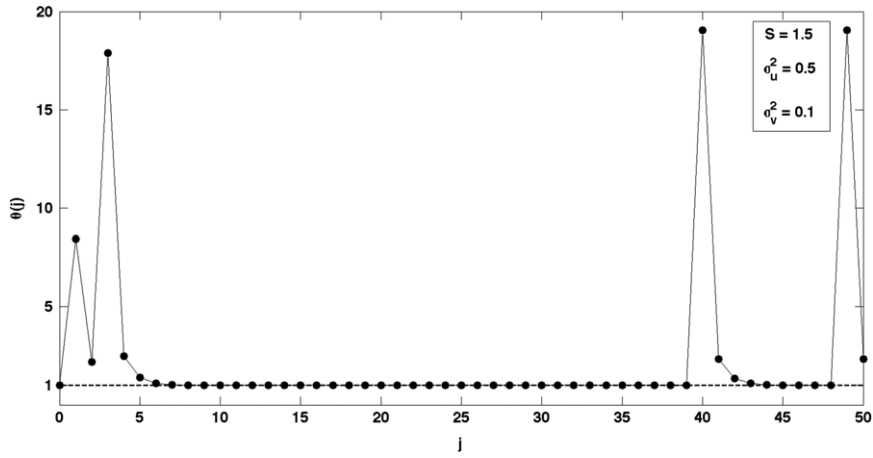
$$P_k(T > r) = P_k(\omega(k+1, k) > S(k+1), \dots, \omega(k+r, k+r-1) > S(k+r)), \quad r = 2, 3, \dots \quad (3.6)$$

Therefore:

$$P_k(T > r) = P_k(T > 1) \cdot \prod_{j=k+1}^{k+r-1} P_k(\omega(j+1, j) > S(j+1) | \omega(k+1, k) > S(k+1), \dots, \omega(j, j-1) > S(j)), \quad r = 2, 3, \dots \quad (3.7)$$

Since $\omega(j+1, j) > S(j+1) \Leftrightarrow \|\check{x}(j+1)\|^2 < \frac{\text{Tr}\{\Psi_\alpha(j+1)\}}{S(j+1)-1}$ then, by introducing the notation $s(j+1) = \frac{\text{Tr}\{\Psi_\alpha(j+1)\}}{S(j+1)-1}$, the generic factor of (3.7) can be written as:

$$\begin{aligned} &P_k(\omega(j+1, j) > S(j+1) | \omega(k+1, k) > S(k+1), \dots, \omega(j, j-1) > S(j)) \\ &= P_k(\|\check{x}(j+1)\|^2 < s(j+1) | \|\check{x}(k+1)\|^2 < s(k+1), \dots, \|\check{x}(j)\|^2 < s(j)) \end{aligned}$$

(a) Measurement saving – $\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio – $\theta(j) = \frac{\text{Tr}\{\Psi_k(j)\}}{\text{Tr}\{\Psi_{\varepsilon}^{(k)}(j)\}}$.**Fig. 7.** Effectiveness of the proposed procedure.

$$\begin{aligned}
 &= P_k(\|A(j)\hat{x}(j)\|^2 < s(j+1) \mid \|\check{x}(k+1)\|^2 < s(k+1), \dots, \|\check{x}(j)\|^2 < s(j)) \\
 &= P_k(\|A(j)\check{x}(j) + A(j)F(j)\rho(j)\|^2 < s(j+1) \mid \|\check{x}(k+1)\|^2 < s(k+1), \dots, \|\check{x}(j)\|^2 < s(j)).
 \end{aligned} \quad (3.8)$$

By defining the following events:

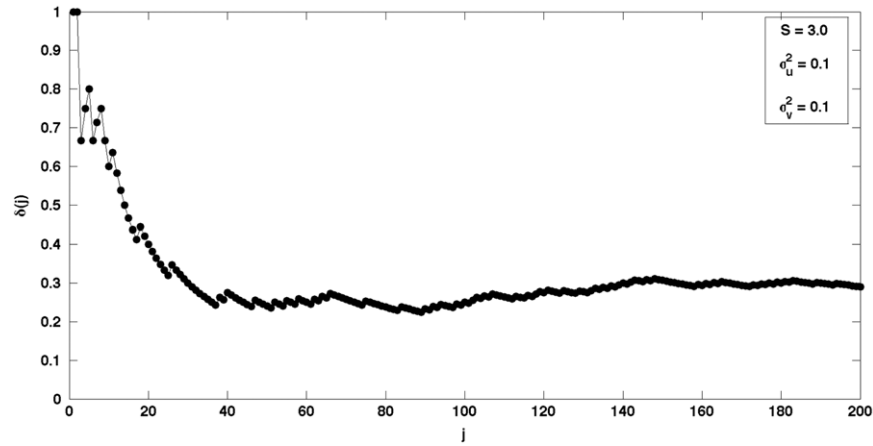
$$\begin{cases}
 \mathcal{A} = \{\|A(j)\check{x}(j) + A(j)F(j)\rho(j)\|^2 < s(j+1)\}, \\
 \mathcal{B} = \{\|\check{x}(j)\|^2 < s(j)\}, \\
 \mathcal{C} = \{\|\check{x}(k+1)\|^2 < s(k+1), \dots, \|\check{x}(j-1)\|^2 < s(j-1)\},
 \end{cases} \quad (3.9)$$

the probability (3.8) can be written as:

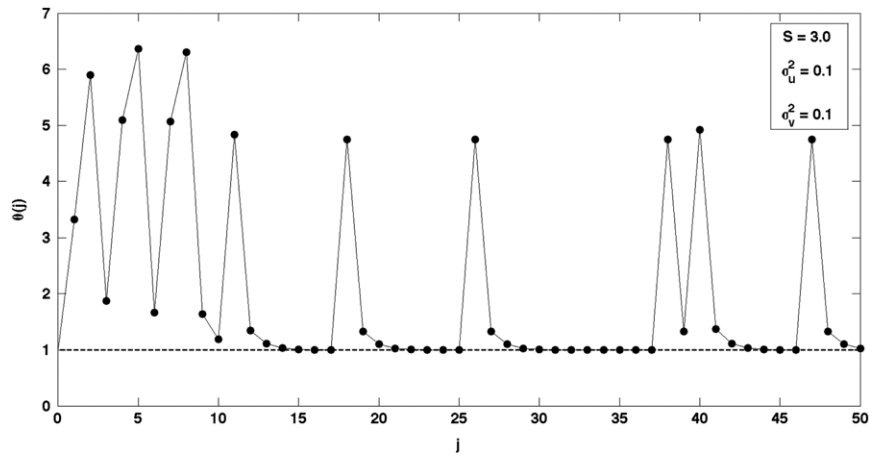
$$P_k(\mathcal{A} \mid \mathcal{B}, \mathcal{C}) = \frac{P_k(\mathcal{A} \cap \mathcal{B} \mid \mathcal{C})}{P_k(\mathcal{B} \mid \mathcal{C})}. \quad (3.10)$$

Since the event \mathcal{B} can be also written as

$$\mathcal{B} = \bigcup_{\xi \in B} \{\|\check{x}(j)\|^2 = \xi\}, \quad \text{where } B = \{\xi \mid \xi \in [0, s(j))\} \quad (3.11)$$



(a) Measurement saving – $\delta(j) = \frac{j-N(j)}{j}$.



(b) Inaccuracy ratio – $\theta(j) = \frac{\text{Tr}(\Psi_{\xi}(j))}{\text{Tr}(\Psi_{\xi}^{(K)}(j))}$.

Fig. 8. Effectiveness of the proposed procedure.

then

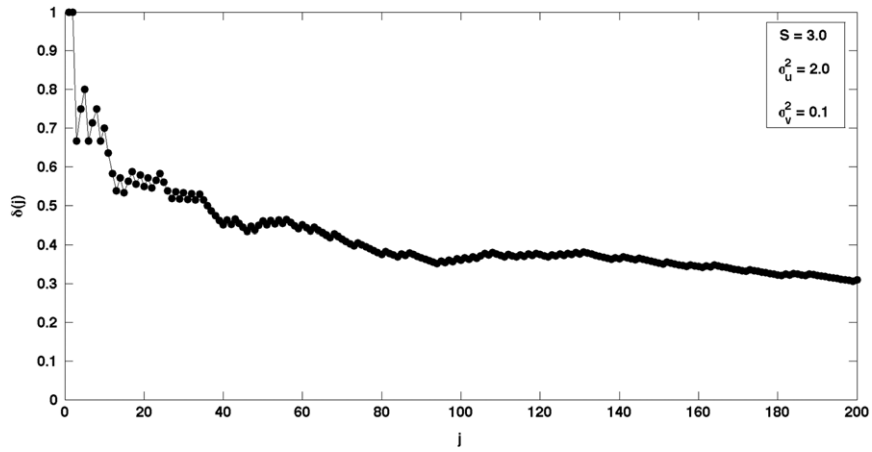
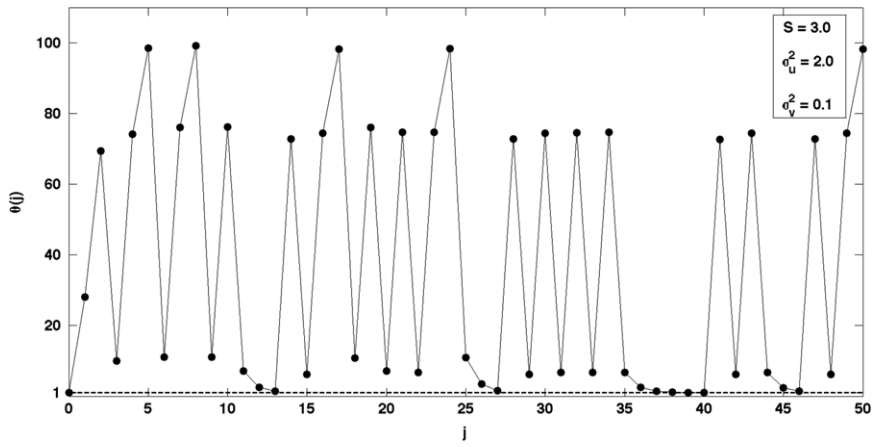
$$\begin{aligned} \frac{P_k(\mathcal{A} \cap \mathcal{B} | \mathcal{C})}{P_k(\mathcal{B} | \mathcal{C})} &= \frac{P_k\left(\mathcal{A} \cap \left[\bigcup_{\xi \in \mathcal{B}} \{\|\check{x}(j)\|^2 = \xi\}\right] | \mathcal{C}\right)}{P_k(\mathcal{B} | \mathcal{C})} = \frac{P_k\left(\bigcup_{\xi \in \mathcal{B}} [\mathcal{A} \cap \{\|\check{x}(j)\|^2 = \xi\}] | \mathcal{C}\right)}{P_k(\mathcal{B} | \mathcal{C})} \\ &= \frac{\sum_{\xi \in \mathcal{B}} P_k(\mathcal{A} \cap \{\|\check{x}(j)\|^2 = \xi\} | \mathcal{C})}{P_k(\mathcal{B} | \mathcal{C})} = \frac{\sum_{\xi \in \mathcal{B}} P_k(\mathcal{A} | \{\|\check{x}(j)\|^2 = \xi\}, \mathcal{C}) P_k(\{\|\check{x}(j)\|^2 = \xi\} | \mathcal{C})}{P_k(\mathcal{B} | \mathcal{C})}. \end{aligned} \quad (3.12)$$

Since $\|\check{x}(j)\|^2$ is a Markov process then the conditioning with respect to \mathcal{C} in $P_k(\mathcal{A} | \{\|\check{x}(j)\|^2 = \xi\}, \mathcal{C})$ can be neglected. So:

$$P_k(\mathcal{A} | \mathcal{B}, \mathcal{C}) = \frac{\sum_{\xi \in \mathcal{B}} P_k(\mathcal{A} | \{\|\check{x}(j)\|^2 = \xi\}) P_k(\{\|\check{x}(j)\|^2 = \xi\} | \mathcal{C})}{P_k(\mathcal{B} | \mathcal{C})}. \quad (3.13)$$

It is immediate to verify that the probability (3.13) can be bounded in this way:

$$P_k(\mathcal{A} | \mathcal{B}, \mathcal{C}) \leq \frac{P_k(\mathcal{A} | \{\|\check{x}(j)\|^2 = 0\}) \sum_{\xi \in \mathcal{B}} P_k(\{\|\check{x}(j)\|^2 = \xi\} | \mathcal{C})}{P_k(\mathcal{B} | \mathcal{C})} \quad (3.14)$$

(a) Measurement saving $-\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio $-\theta(j) = \frac{\text{Tr}\{\Psi_{\varepsilon}(j)\}}{\text{Tr}\{\Psi_{\varepsilon}^{(K)}(j)\}}$.**Fig. 9.** Effectiveness of the proposed procedure.

since

$$P_k(\mathcal{A}|\{\|\check{x}(j)\|^2 = \xi\}) \leq P_k(\mathcal{A}|\|\check{x}(j)\|^2 = 0).$$

Then, recalling (3.11), we have

$$P_k(\mathcal{A}|\mathcal{B}, \mathcal{C}) \leq P_k(\mathcal{A}|\|\check{x}(j)\|^2 = 0). \quad (3.15)$$

Denoting

$$\begin{aligned} \gamma(j+1) &= P_k(\mathcal{A}|\{\|\check{x}(j)\|^2 = 0\}) \\ &= P_k(\|A(j)\check{x}(j) + A(j)F(j)\rho(j)\|^2 < s(j+1)\|\check{x}(j)\|^2 = 0) \\ &= P_k(\|A(j)F(j)\rho(j)\|^2 < s(j+1)), \quad j = k+1, k+2, \dots, k+r-1, \end{aligned} \quad (3.16)$$

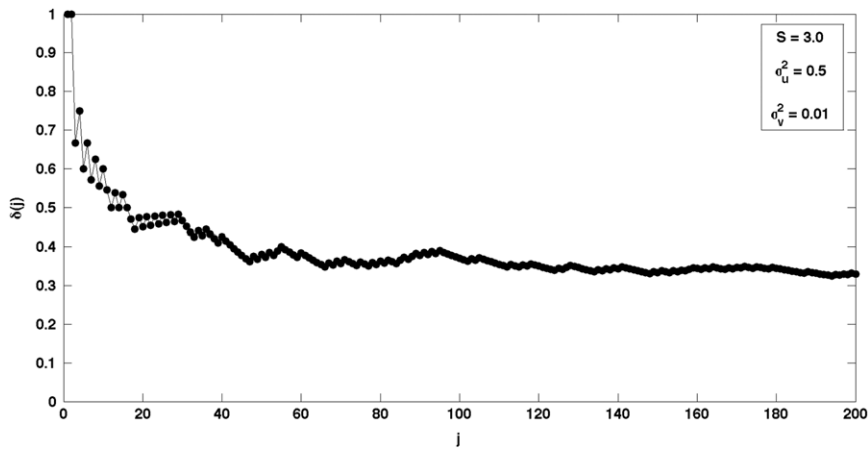
the probability (3.6) can be bounded in this way:

$$P_k(T > r) \leq P_k(T > 1) \prod_{j=k+1}^{k+r-1} \gamma(j+1), \quad r = 2, 3, \dots \quad (3.17)$$

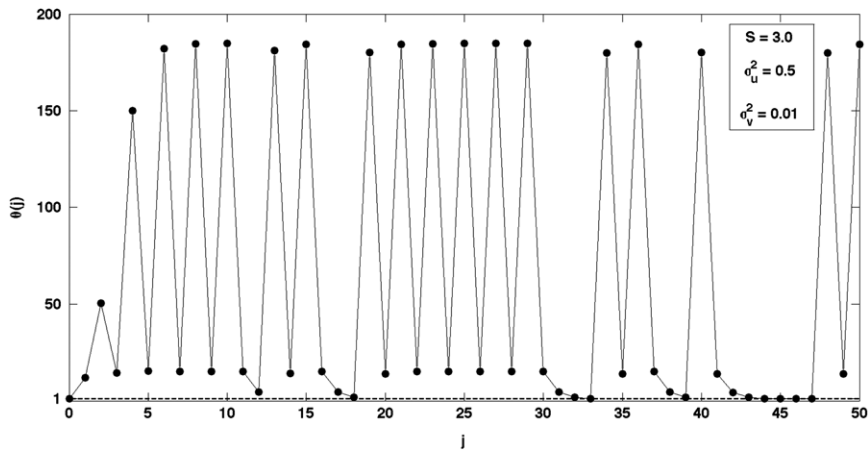
Recalling the well known equality

$$E_k\{T\} = \sum_{r=0}^{\infty} P_k(T > r) = 1 + \sum_{r=1}^{\infty} P_k(T > r) \quad (3.18)$$

the inequality (3.3) immediately follows. \square



(a) Measurement saving – $\delta(j) = \frac{j-N(j)}{j}$.



(b) Inaccuracy ratio – $\theta(j) = \frac{\text{Tr}\{\psi_\varepsilon(j)\}}{\text{Tr}\{\psi_\varepsilon^{(k)}(j)\}}$.

Fig. 10. Effectiveness of the proposed procedure.

From Theorem 3.1 the next result immediately follows.

Corollary 3.2. *The condition*

$$\sum_{r=2}^{\infty} P_k(T > r) \prod_{j=k+1}^{k+r-1} \gamma(j+1) < \infty \quad (3.19)$$

is sufficient to establish that:

$$P_k(T < \infty) = 1, \quad (3.20)$$

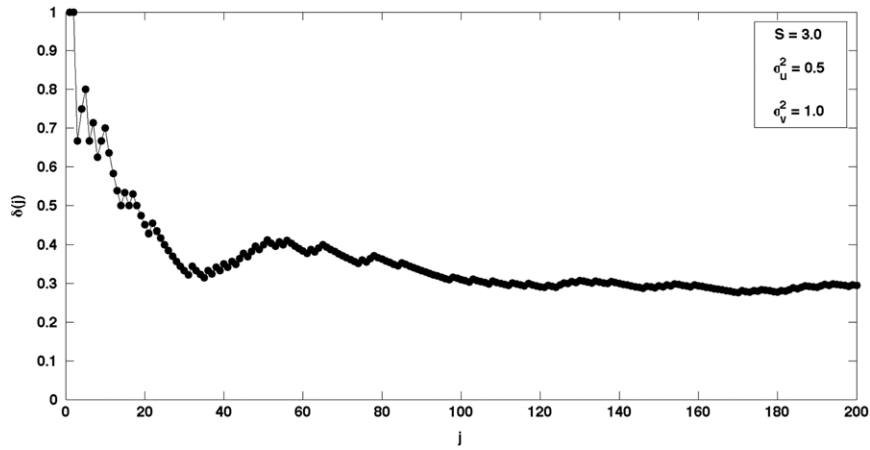
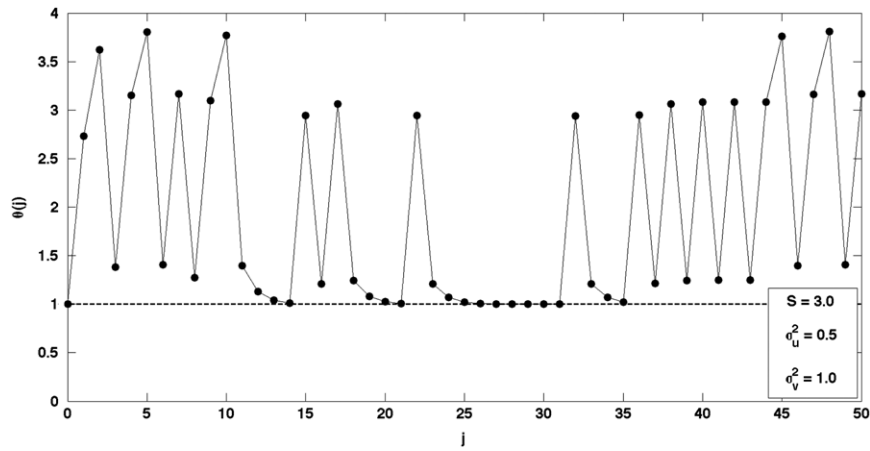
as well as:

$$E_k\{T\} < \infty. \quad (3.21)$$

Proof. As is well known, the Borel–Cantelli lemma provides:

$$\sum_{r=0}^{\infty} P_k(T > r) < \infty \quad (3.22)$$

as a sufficient condition for (3.20) to hold. Recalling (3.18), we see that condition (3.22) is sufficient for (3.21) as well. \square

(a) Measurement saving $-\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio $-\theta(j) = \frac{\text{Tr}\{\psi_e(j)\}}{\text{Tr}\{\psi_e^{(K)}(j)\}}$.**Fig. 11.** Effectiveness of the proposed procedure.

A looser upper bound for $P_k(T > r)$, which however is easier to be computed is provided by the following result.

Theorem 3.3. For $r = 2, 3, \dots$, we have that

$$P_k(T > r) \leq P_k(T > 1) \prod_{j=k+1}^{k+r-1} \beta(j+1), \quad (3.23)$$

where

$$\beta(j+1) = \int_0^{\frac{s(j+1)}{\lambda_m(j)}} \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx \quad (3.24)$$

and $\lambda_m(j)$ is the smallest eigenvalue of the square symmetric positive semidefinite matrix

$$Q(j) = \Psi_\rho^{T\frac{1}{2}}(j) F^T(j) A^T(j) A(j) F(j) \Psi_\rho^{\frac{1}{2}}(j). \quad (3.25)$$

Proof. The innovation process $\rho(j)$ can be written as

$$\rho(j) = \Psi_\rho^{\frac{1}{2}}(j) \xi(j) \quad (3.26)$$

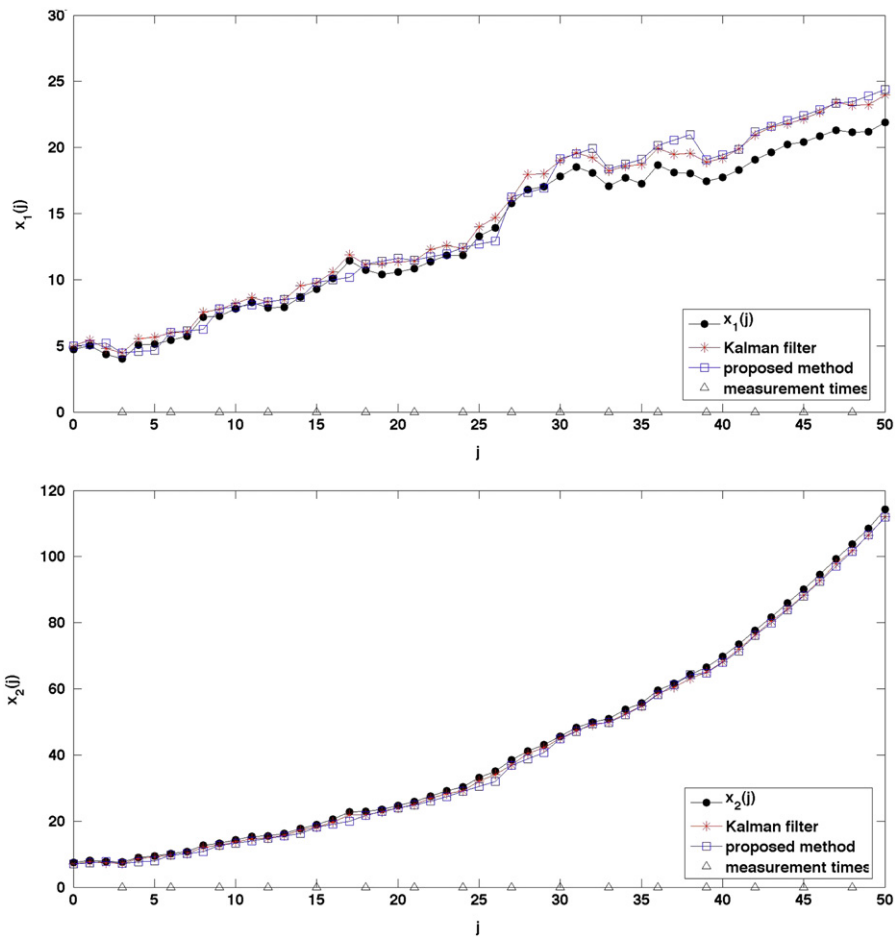


Fig. 12. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

where the vector $\xi(j)$ is a Gaussian vector with zero mean and covariance matrix $\Psi_\xi(j) = I$, since $\rho(j)$ is a Gaussian vector with zero mean and a known covariance matrix

$$\Psi_\rho(j) = C(j)[A(j-1)\Psi_\epsilon(j-1)A^T(j-1) + B(j-1)\Psi_u(j-1)B^T(j-1)]C^T(j) + \Psi_v(j). \quad (3.27)$$

Then (3.4) can be written as:

$$\gamma(j+1) = P_k(\xi^T(j)Q(j)\xi(j) < s(j+1)), \quad j = k+1, \dots, k+r-1. \quad (3.28)$$

Denoting by $\Lambda(j)$ a diagonal matrix that has the same eigenvalues of $Q(j)$ with the corresponding multiplicity, it results:

$$Q(j) = T^T(j)\Lambda(j)T(j), \quad (3.29)$$

where $T(j)$ is an orthogonal matrix [18]. Using Eq. (3.29) the probability (3.28) can be written as

$$\gamma(j+1) = P_k(\xi^T(j)T^T(j)\Lambda(j)T(j)\xi(j) < s(j+1)) = P_k(\eta^T(j)\Lambda(j)\eta(j) < s(j+1)), \quad (3.30)$$

where the vector $\eta(j) = T(j)\xi(j)$ is a standard Gaussian vector.

Denoting by $\lambda_i(j)$, $i = 1, 2, \dots, n$, the (real non-negative) eigenvalues of $Q(j)$, then the probability (3.30) can be written as

$$\gamma(j+1) = P_k\left(\sum_{i=1}^n \lambda_i(j)\eta_i^2(j) < s(j+1)\right), \quad (3.31)$$

where $\eta_i(j)$, $i = 1, 2, \dots, n$, are the components of $\eta(j)$. Since

$$\sum_{i=1}^n \lambda_i(j)\eta_i^2(j) \geq \lambda_m(j) \sum_{i=1}^n \eta_i^2(j), \quad (3.32)$$

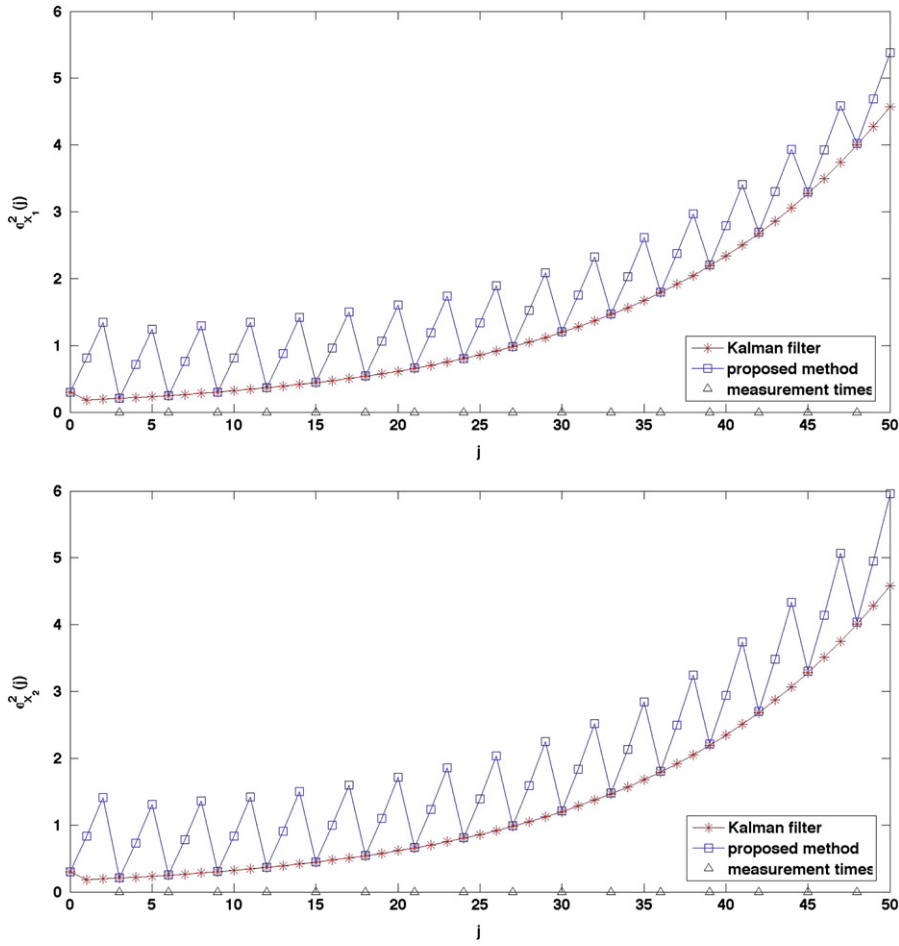


Fig. 13. Evolution of covariance matrix elements $\sigma_{x_1}^2(j)$ and $\sigma_{x_2}^2(j)$.

where $\lambda_m(j)$ is the smallest eigenvalue of $\Gamma(j)$, then we have:

$$\gamma(j+1) \leq P_k \left(\lambda_m(j) \sum_{i=1}^n \eta_i^2(j) < s(j+1) \right) = \beta(j+1) \quad (3.33)$$

from which, taking into account the inequality (3.2) of Theorem 3.1, the relation (3.23) immediately follows. Assuming that $\lambda_m(j) > 0$, the probability $\beta(j+1)$ can be written as

$$\beta(j+1) = P_k \left(\sum_{i=1}^n \eta_i^2(j) \leq \frac{s(j+1)}{\lambda_m(j)} \right) \quad (3.34)$$

and the last equality (3.24) of Theorem 3.3 follows from the fact that the random variable $Y = \sum_{i=1}^n \eta_i^2(j)$ is the sum of squares of standard Gaussian variables $\eta_i(j)$ and therefore has a χ^2 distribution with n degrees of freedom.

We can notice that in the case of $\lambda_m(j) = 0$ the results of Theorem 3.3 are also valid but relation (3.23) does not give a significative upper bound for $P_k(T > r)$ since, according to (3.33), we have $\beta(j+1) = 1$. \square

From Theorem 3.3 the following result immediately follows.

Corollary 3.4. *If*

$$\frac{s(j+1)}{\lambda_m(j)} \leq c_1, \quad j = k+1, k+2, \dots, \quad (3.35)$$

for a finite positive constant c_1 , then:

$$P_k(T > r) \leq c_2^{(r-1)}, \quad r = 2, 3, \dots, \quad (3.36)$$

and

$$E_k\{T\} \leq \frac{2 - c_2}{1 - c_2}, \quad (3.37)$$

with

$$c_2 = \int_0^{c_1} \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx < 1. \quad (3.38)$$

Proof. The proof is immediate from (3.35) since then:

$$\beta(j+1) \leq \int_0^{c_1} \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} dx = c_2. \quad \square \quad (3.39)$$

Remark 3.5. All the results in this section do hold whether k is a measurement time or not. In any case $T + k$ keeps the meaning of the first non-measurement time after k .

4. How long to keep on not measuring the output process?

Let k be a time in which we performed an output process measurement ($k = 0$ if no measurement occurred yet) and let us assume that we did not measure $z(k+1)$. Then, taking into account the procedure based on the function (2.16) described in Section 2, the “non-measurement” interval T' will be such that the function $\omega(j, k)$ stays below or equal to the given threshold $S(j)$, $j = k+1, \dots, k+T'-1$, while $\omega(k+T', k)$ exceeds $S(k+T')$.

The question now is whether, and under which conditions, such a T' exists finite. For this purpose we first note that due to the very definition (2.16) the function $\omega(j, k)$ is monotone increasing w. r. t. j and in particular:

$$\omega(j, k) \geq (j - k). \quad (4.1)$$

Therefore, as soon as $S(j)$ enjoys a (strictly) sublinear growth, a (first) finite T' exists such that:

$$\omega(k+T', k) > S(k+T') \quad (4.2)$$

and we resume measuring z at $k+T'$.

In particular, for a constant threshold $S(j) = S$, (4.1) and (4.2) immediately yield the upper bound for T' :

$$T' \leq [S] + 1, \quad (4.3)$$

with $[S]$ the integer part of S .

Remark 4.1. Note that the closer $\chi(i)$ in (2.16) is to its lower bound 1, the closer T' is to its upper bound (4.3), thus being well approximated by the largest integer smaller than (4.3) itself, that is $[S]$. Under the same circumstances it follows from (2.16) that $\omega(j, k) \approx j - k$. Then for $S > 1$ we get $T = 1$ as well as a periodic measurement/non-measurement switching corresponding to $T = 1$, $T' = [S]$. This in particular happens whenever $\|\check{x}(i)\|^2$ diverges and at the same time $\text{Tr}\{\Psi_\alpha(i)\}$ goes to a limit, such as in unstable (for the first property) and in controllable and observable (for the second property) time-invariant systems.

A remark similar to Remark 3.5 still applies to the results in this section.

5. Evaluation of the proposed procedure against the standard Kalman filter

5.1. A stable system

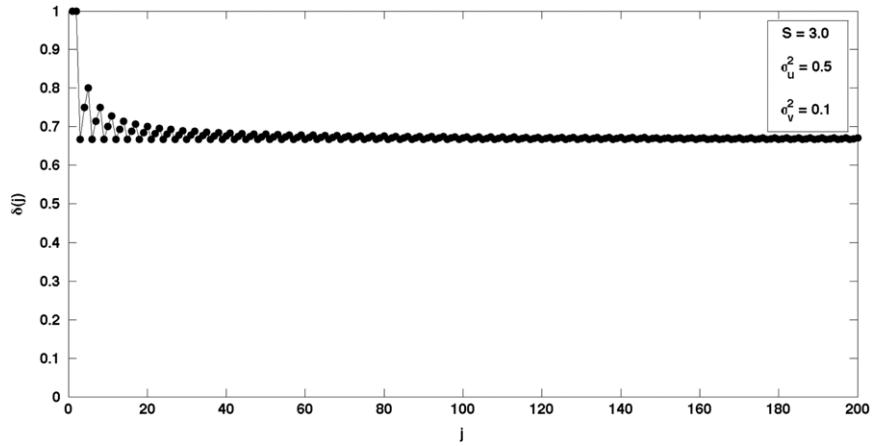
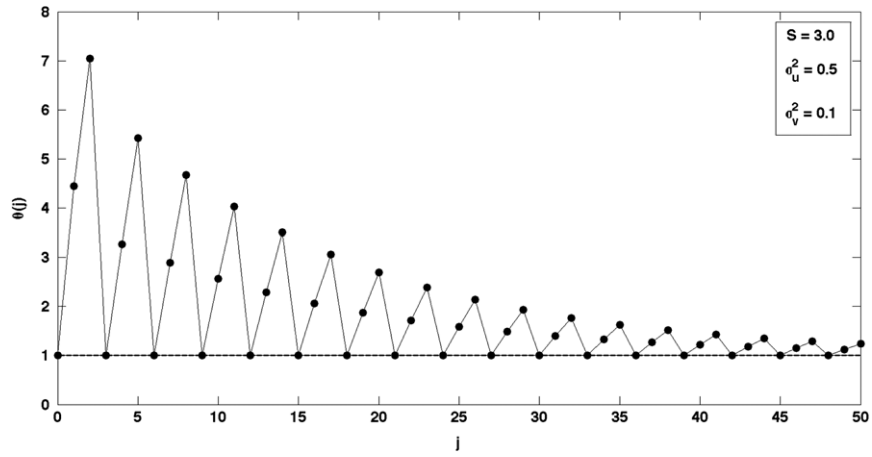
Let us now consider the following linear discrete time-invariant dynamical system:

$$\begin{cases} x(j+1) = Ax(j) + Bu(j), & j = 0, 1, 2, \dots, \\ z(j) = Cx(j) + v(j), & j = 1, 2, \dots, \end{cases} \quad (5.1)$$

where the state vector $x(j) \in \mathbb{R}^2$ and we assumed:

$$A = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.8 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \end{bmatrix}. \quad (5.2)$$

Let $x(0) \in \mathbb{R}^2$ be a stochastic Gaussian vector and let $\{u(j)\}$, $\{v(j)\}$ be white Gaussian noises, independent of each other and of $x(0)$ as well.

(a) Measurement saving – $\delta(j) = \frac{j-N(j)}{j}$.(b) Inaccuracy ratio – $\theta(j) = \frac{\text{Tr}[\psi_e(j)]}{\text{Tr}[\psi_e^{(k)}(j)]}$.**Fig. 14.** Effectiveness of the proposed procedure.

The main purpose of this section is to evaluate the effectiveness of [Decision Rule 2.1](#) proposed in [Section 2](#) by considering some simulations of the dynamical system [\(5.1\)–\(5.2\)](#) and by estimating the state vector both with the procedure described by the above decision rule and with the Kalman filter. For any simulation we evaluated the saving of measurements obtained with the proposed procedure, by computing the ratio between the number of measurements actually performed and the total number of measurements processed by the Kalman filter. We also evaluated the downgrading of the estimate accuracy (produced by the non-measurement times) compared with the Kalman filter accuracy, by computing the ratio between the estimate mean square error of the proposed procedure and the estimate mean square error of the Kalman filter itself.

Let us first consider a simulation obtained by choosing the following mean values and covariances for $x(0)$, $\{u(j)\}$, $\{v(j)\}$:

$$\bar{x}(0) = \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \quad \Psi_x(0) = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad (5.3)$$

$$\bar{u}(j) = 0, \quad \sigma_u^2(j) = 0.5, \quad j = 0, 1, 2, \dots, \quad (5.4)$$

$$\bar{v}(j) = 0, \quad \sigma_v^2(j) = 0.1, \quad j = 1, 2, \dots, \quad (5.5)$$

and a constant threshold:

$$S = 3.0. \quad (5.6)$$

In [Fig. 1](#) the evolution of $x_1(j)$ and $x_2(j)$ is represented, along with both estimates of them obtained following the two different procedures. In [Fig. 2](#) the evolution of covariance matrix diagonal elements of both estimates, respectively $\sigma_{x_1}^2(j)$ and $\sigma_{x_2}^2(j)$ is shown.

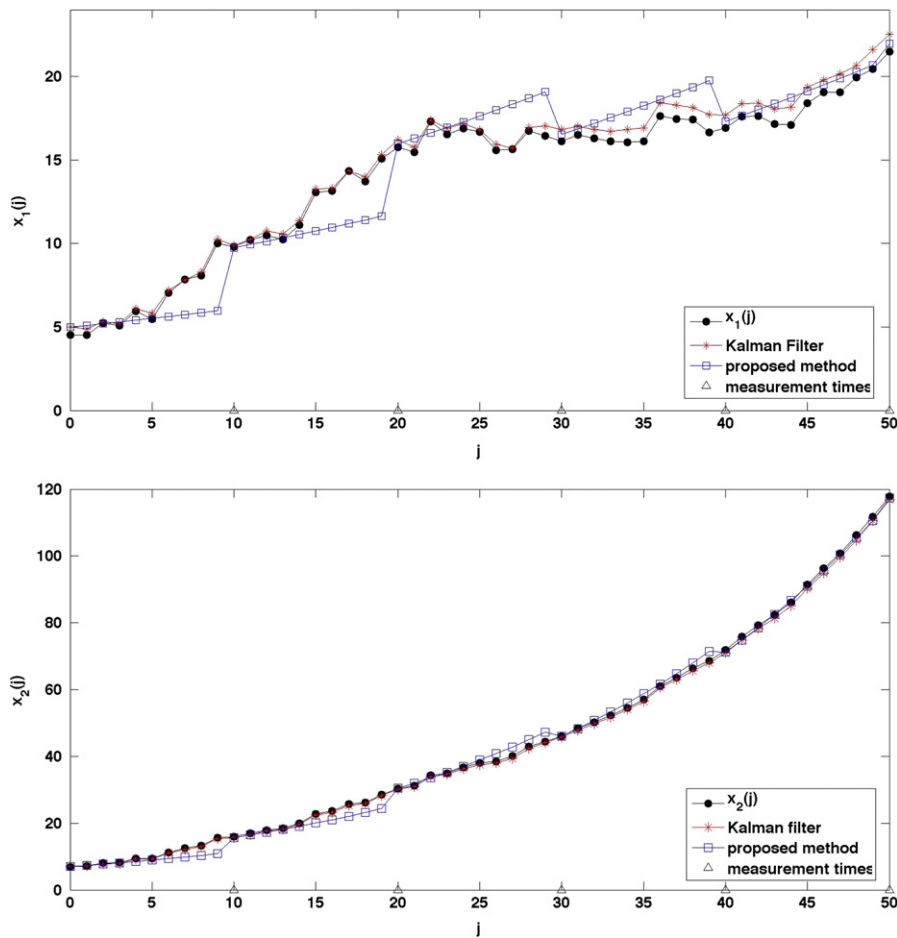


Fig. 15. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

Let us now consider the relative saving of measurements obtained with the proposed procedure. In Fig. 3(a) the evolution of the ratio $\delta(j) = \frac{j-N(j)}{j}$, where $N(j)$ is the number of the measurements actually processed by the proposed procedure up to time j is represented. It is easy to notice that this evolution asymptotically goes to a value, about 0.3. Therefore, in this case, the procedure allows a saving in terms of measurement number of about 30%.

Nevertheless, this measurement saving implies a downgrading of the estimate accuracy. This downgrading is clearly pointed out in Fig. 3(b) in which the evolution of the inaccuracy ratio $\theta(j)$ between the mean square error of state estimate using our procedure, $\text{Tr}\{\Psi_e(j)\}$, and the mean square error of the Kalman filter, $\text{Tr}\{\Psi_e^{(K)}(j)\}$ is represented. The ratio is considerably higher than 1.0 correspondingly to the non-measurement times, but this downgrading is rapidly recovered when a new measurement is made and the ratio quickly goes down to 1.0.

Let us now consider some simulations obtained with the same stochastic characterization of $x(0)$, $\{u(j)\}$, $\{v(j)\}$ given by (5.3)–(5.5) and by varying the threshold S . Let us first consider a higher threshold $S = 100$.

In Fig. 4 the evolution of $x_1(j)$, $x_2(j)$ is represented along with both estimates of them obtained by following the two different procedures.

We notice that now measurement times occur more rarely and consequently the saving of measurements reaches 85% of the total available amount (Fig. 5(a)). On the other hand $\theta(j)$ periodically swings from 1.0 to an asymptotic value of about 35 (corresponding to the ratio between the asymptotic prediction and filtering total error estimate covariances), as is shown in Fig. 5(b).

If we consider a lower threshold $S = 1.5$ the achieved results are shown in Figs. 6 and 7, with an asymptotic measurement saving of about 10% and an inaccuracy ratio equal to 1.0 for most of the time.

Let us now see what happens if we change the values of σ_u^2 and σ_v^2 , coming back to the value $S = 3.0$. From Figs. 8 and 9 it is simple to notice that $\delta(j)$ remains asymptotically unchanged (equal to 0.3) against the input variance σ_u^2 , while $\theta(j)$ increases when σ_u^2 increases.

Similarly $\delta(j)$ remains asymptotically unchanged (equal to 0.3) against the output variance σ_v^2 , while $\theta(j)$ decreases when σ_v^2 increases, as is shown in Figs. 10 and 11.

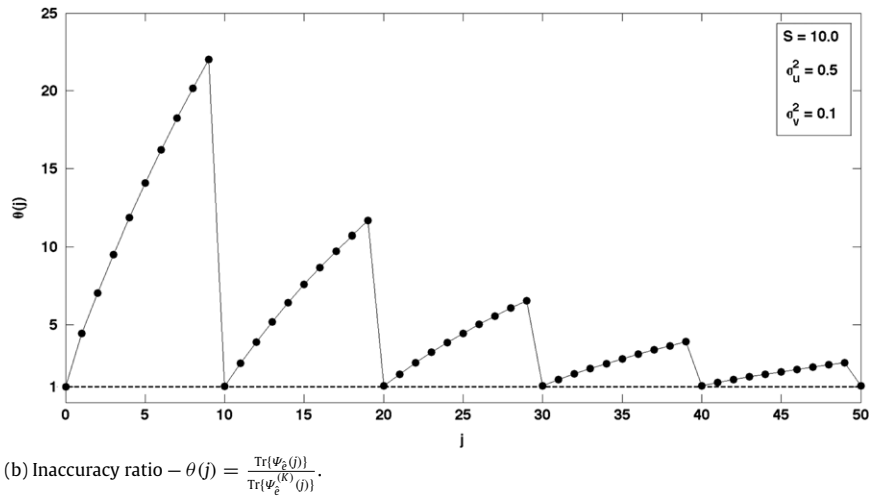
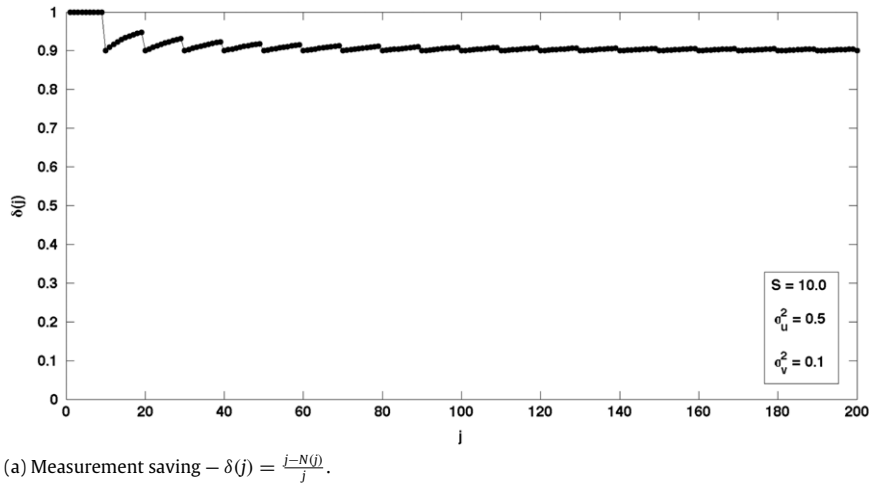


Fig. 16. Effectiveness of the proposed procedure.

5.2. An unstable system

Let us now change the dynamics of system (5.1) and let us assume:

$$A = \begin{bmatrix} 1.02 & 0 \\ 0 & 1.05 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \end{bmatrix}. \quad (5.7)$$

System (5.1)–(5.7) is now unstable. Let us take the values (5.3), (5.4), (5.5) for the mean values and covariances of $x(0)$, $\{u(j)\}$, $\{v(j)\}$ and the value (5.6) for the threshold S .

In Fig. 12 the evolution of $x_1(j)$ and $x_2(j)$ is represented along with the corresponding estimates. In Fig. 13 the covariance matrix diagonal elements $\sigma_{x_1}^2(j)$, $\sigma_{x_2}^2(j)$ of both estimates are represented. It can be noticed from the figures that, in this case, the proposed procedure suggests to measure with a great regularity (a measurement at each third time instant), according to the content of Remark 4.1. In Fig. 14(a) the saving of measurements obtained with the proposed procedure is also represented. This evolution asymptotically goes to a value, about 0.67. Therefore measurement saving in the case of system (5.1) and (5.7) is higher than in the previous example, that is in system (5.1) and (5.2). Moreover the inaccuracy ratio of this case is lower than that in the previous case and its value quickly decreases with time, as is shown in Fig. 14(b).

Let us now consider the same system (5.1) and (5.7) with the mean values and covariances (5.3) and (5.4), (5.5) and with a higher threshold, $S = 10$.

In this case too, as shown in Fig. 15, the proposed procedure suggests to measure with regularity but now the non-measurement interval T' is equal to 10, according to Remark 4.1. The measurement saving is about 90% (Fig. 16(a)). The

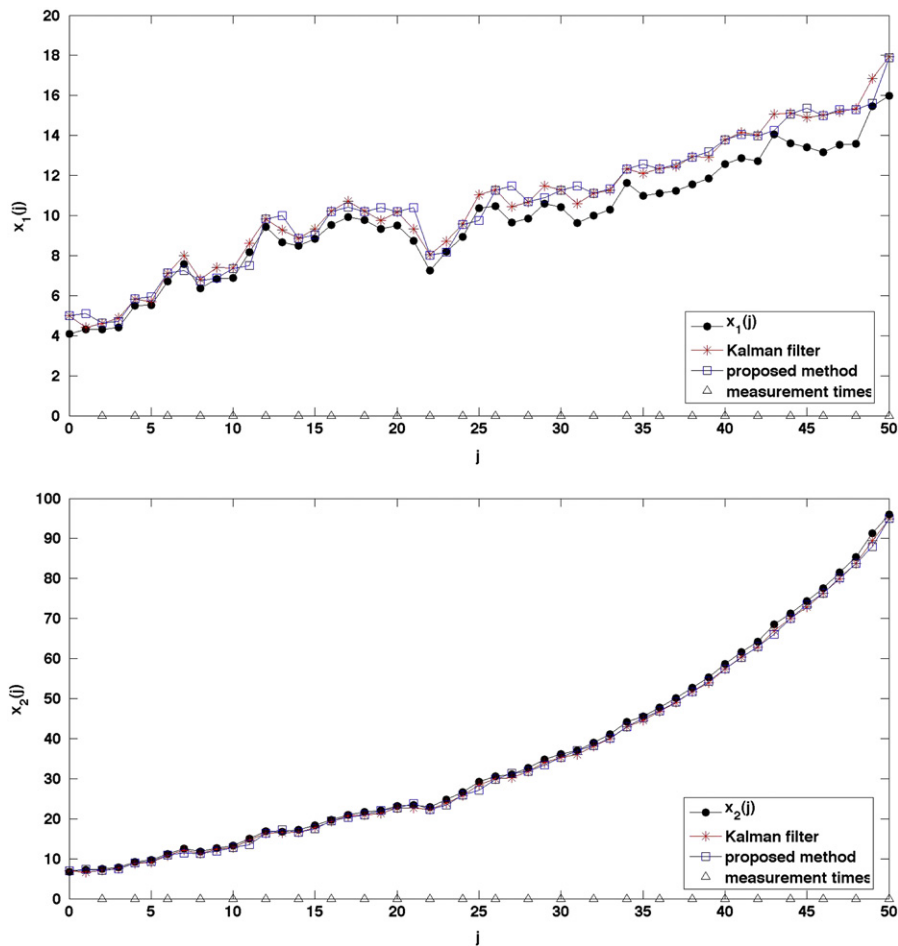


Fig. 17. Evolution of $x_1(j)$ and $x_2(j)$ along with their estimates.

estimate accuracy obviously decreases during the non-measurement interval T' , as shown in Fig. 16(b) where the inaccuracy ratio increases between two measurement times.

Let us now reduce the threshold down to 1.5 for system (5.1) and (5.7) without changing the mean values and covariances (5.3)–(5.5) (Fig. 17).

We notice that now measurement saving reaches only 50% of the total available amount (Fig. 18(a)), lower than 67% value of the case of Fig. 14.

The results obtained by varying the input and output variances in the case of system (5.1), (5.2) can be confirmed also in case of system (5.1), (5.7): $\delta(j)$ remains unchanged against the input and measurement noise variances, while $\theta(j)$ increases when σ_u^2 increases and decreases when σ_v^2 increases. The main difference with the stable system example, as it has been noticed above, is that $\theta(j)$ takes lower values, decreasing with time.

6. Concluding remarks

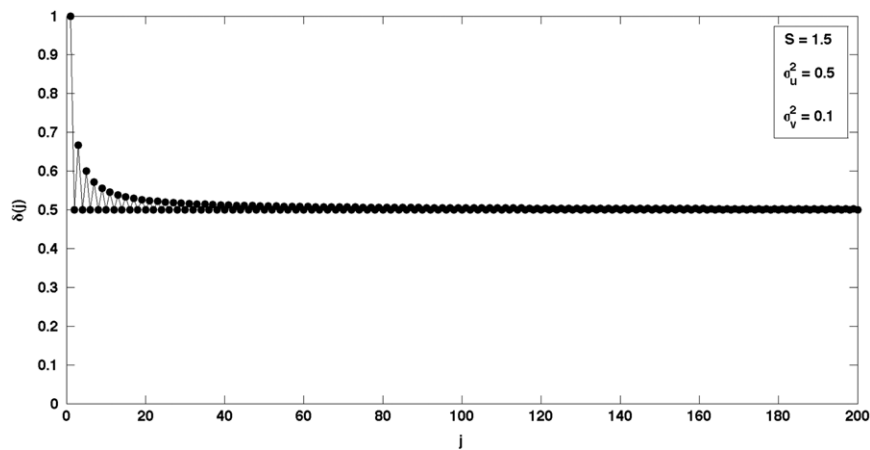
In this paper we discussed the problem of deciding when to actually measure and process the output values, in order to perform the estimate of the state of a (discrete time) linear system.

We showed that it is possible to design a convenient measurement policy, able to reduce the cost of the measurement itself, while keeping the estimate inaccuracy within acceptable values.

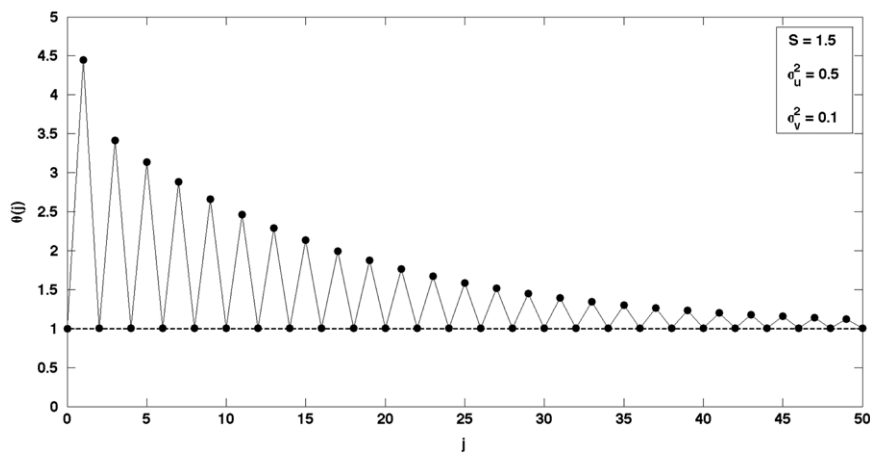
The relevance of a measurement policy is especially acknowledged in those less traditional fields where the measurement cost is not only economically assessed but is related to possible damages of the measured “object”.

We analysed by some probabilistic arguments the upper bounds for the measurement and non-measurement time intervals.

We also investigated in a few examples the influence of the decision threshold (measurement cost) as well as that of the state and output noise variances over the outcome of the proposed procedure.



(a) Measurement saving $-\delta(j) = \frac{j-N(j)}{j}$.



(b) Inaccuracy ratio $-\theta(j) = \frac{\text{Tr}\{\Psi_e(j)\}}{\text{Tr}\{\Psi_e^{(K)}(j)\}}$.

Fig. 18. Effectiveness of the proposed procedure.

References

- [1] N. Wiener, *Cybernetics: or Control and Communication in the Animal and the Machine*, Mass. and Wiley, New York, 1948.
- [2] S. Tatikonda, S. Mitter, Control under communication constraints, *IEEE Transactions on Automatic Control* 49 (7) (2004) 1056–1068.
- [3] N. Elia, When Bode meets Shannon: Control-oriented feedback communication schemes, *IEEE Transactions on Automatic Control* 49 (9) (2004) 1477–1488.
- [4] J. Liu, N. Elia, Convergence of fundamental limitations in information, estimation and control, in: *Proceedings of the 45th IEEE Conference on Decision & Control*, San Diego, December 2006, pp. 5609–5614.
- [5] P. Padmasola, N. Elia, Bode integral limitations of spatially invariant multi-agent systems, in: *Proceedings of the 45th IEEE Conference on Decision & Control*, San Diego, December 2006, pp. 4327–4332.
- [6] S. Tanaka, T. Okita, Optimal timing of observations for state estimation in a one-dimensional linear continuous system, *Automatica* 21 (3) (1985) 329–331.
- [7] H.J. Kushner, On the optimum timing of observations for linear control systems with unknown initial state, *IEEE Transactions on Automatic Control* 9 (2) (1964) 144–150.
- [8] M. Aoki, M.T. Li, Optimal discrete-time control system with cost for observation, *IEEE Transactions on Automatic Control* 14 (2) (1969) 165–175.
- [9] B.M. Miller, W.J. Runggaldier, Optimization of observations: A stochastic control approach, *SIAM Journal on Control and Optimization* 35 (3) (1997) 1030–1052.
- [10] A.A. Fasolya, V.V. Khutortsev, Local optimization of measuring processes, *Automation and Remote Control* 63 (8) (2002) 1279–1284.
- [11] D.J. Mellefont, R.W.H. Sargent, Optimal measurement policies for control purpose, *International Journal of Control* 26 (4) (1977) 595–602.
- [12] A. Sano, M. Terao, Measurement optimization in optimal process control, *Automatica* 6 (5) (1970) 705–714.
- [13] Y. Sawaragi, T. Soeda, Y. Tomita, I. Imai, Optimal timing of the observations for the state estimation and control of the stochastic discrete linear system, *International Journal of Control* 27 (4) (1978) 621–637.
- [14] V.M. Moreno, A. Pigazo, *Kalman Filter: Recent Advances and Applications*, InTech, Croatia, 2009.

- [15] F. Morbidi, A. Garulli, D. Prattichizzo, C. Rizzo, S. Rossi, Application of Kalman filter to remove TMS-induced artifacts from EEG recordings, *IEEE Transactions on Control Systems Technology* 16 (6) (2008) 1360–1366.
- [16] R. Kazemi, A. Farsi, M.H. Ghaed, M. Karimi-Ghartemani, Detection and extraction of periodic noises in audio and biomedical signals using Kalman filter, *Signal Processing* 88 (8) (2008) 2114–2121.
- [17] V.P. Oikonomou, A.T. Tzallas, D.I. Fotiadis, A Kalman filter based methodology for EEG spike enhancement, *Computer Methods and Programs in Biomedicine* 85 (2) (2007) 101–108.
- [18] M.H.A. Davis, R.B. Vinter, *Stochastic Modelling and Control*, Chapman and Hall, New York, 1985.